On the Routh-Hurwitz-Fujiwara and the Schur-Cohn-Fujiwara Theorems for the Root-Separation Problem

B. N. Datta*

Department of Mathematics
Ahmadu Bello University,
Zaria, Nigeria.

Submitted by David Carlson

ABSTRACT

In this paper, we give simple and elementary proofs of the two classical results of Fujiwara on the solution of the well-known Routh-Hurwitz and Schur-Cohn problems. We show that the Fujiwara matrix in each case satisfies a Lyapunov-type equation and then obtain Fujiwara's results by applying to this matrix equation some recent results on the inertia of matrices. These alternative proofs of Fujiwara's results thus establish a link between two apparently different approaches to the solution of the root-separation problem: the classical method of solution via quadratic forms, and the solution via matrix equations.

1. INTRODUCTION

Let $f(x)=x^n-a_nx^{n-1}\ldots a_2x-a_1$ be a given polynomial with real coefficients. Then the classical Routh-Hurwitz problem is the problem of finding the number of zeros of $f(x)$ with negative real parts, and in particular of obtaining a necessary and sufficient condition for all zeros to lie in the left half plane ($\mathrm{Re}\{z\}<0$). The Schur-Cohn problem is the one of determining the number of zeros inside a unit circle and in particular of establishing a necessary and sufficient condition for all the zeros to lie inside it. Different methods of solutions of these problems are available in the literature (see,

*Present Address: Instituto de Matemática, Estatistica e Ciências da Computação, Universidade Estadual de Campinas, Campinas, Brasil.

0024-3795/78/0022-0235$01.75

e.g., the survey [10] of Krein and Naimark). Among these, an elegant one is due to Fujiwara [6], who gave a unified treatment of both problems.

Fujiwara's method of solution can be described as follows: Given $f(x)$, he defined another suitable polynomial $f^*(x)$. From $f(x)$ and $f^*(x)$, he constructed the Bezoutiant bilinear form

$$K(f) = \frac{f(x)f^*(y) - f(y)f^*(x)}{x-y}$$

$$= \sum_{i,k=0}^{n-1} b_{ik} x^i y^k. \tag{1}$$

The matrix $B=(b_{ik})$ is called the Bezout matrix. From it, he formed another symmetric matrix, F, known as the Fujiwara matrix, and obtained the solution of each problem in terms of the inertia of the matrix F. The inertia of a matrix A is defined to be a triplet $(\pi(A),\nu(A),\delta(A))$, where $\pi(A)$, $\nu(A)$ and $\delta(A)$ are, respectively, the numbers of eigenvalues of A with positive, negative and zero real parts. It is denoted by $\mathrm{In}(A)$. The controllability matrix $(N,AN,A^2N,\dots,A^{n-1}N)$ will be denoted by $[A,N]$.

We now state the fundamental results of Fujiwara.

THEOREM 1 (Routh, Hurwitz, Fujiwara). *Let $f^*(x)=f(-x)$, and let the Fujiwara matrix $F=(f_{ik})$ be defined from the Bezout matrix $B=(b_{ik})$ of $f(x)$ and $f^*(x)$ as*

$$f_{ik} = (-1)^i b_{ik}, \qquad i,k = 0,1,\dots,n-1. \tag{2}$$

Assume that F is nonsingular. Then:

(a) *The numbers of zero of $f(x)$ with negative and positive real parts are respectively equal to the numbers of positive and negative eigenvalues of F.*

(b) *F is positive definite if and only if all the zeros of $f(x)$ have negative real parts.*

THEOREM 2 (Schur, Cohn, Fujiwara). *Let $f^*(x)=x^n f(1/x)$ and let the Fujiwara matrix $F=(f_{ik})$ be defined as*

$$f_{ik} = b_{i,n-1-k}, \tag{3}$$

where $B=(b_{ik})$ is the associated Bezout matrix of $f(x)$ and $f^*(x)$. Then:

(a) *If $\pi(F)$ and $\nu(F)$ are the numbers of positive and negative eigenvalues of F, and F is nonsingular $[\pi(F)+\nu(F)=n]$, then $f(x)$ has $\pi(F)$ zeros inside the unit circle and $\nu(F)$ outside it.*

(b) *F is positive definite if and only if all the zeros lie inside the unit circle.*

In this paper, we give proofs of Theorem 1 and Theorem 2 via Lyapunov matrix equations. We shall show that the Fujiwara matrix F in each case satisfies a Lyapunov-type equation, and Fujiwara's result then follows from some recent results on the inertia of matrices. Parks [12] first noted the connection of the Routh-Hurwitz problem with the Lyapunov equation, and then Lehnigk [11] proved Hermite's stability criterion and Hermite's theorem on the number of zeros with positive real parts of a complex polynomial, using Lyapunov's functions. Howland [8] pointed out that Fujiwara matrix of the Routh-Hurwitz problem satisfies a Lyapunov-type equation. Kalman [9] proved part (b) of Theorem 2 using the second method of Lyapunov. However, our proofs are different from those of Lehnigk or Kalman, and they seem simpler. Our alternative proofs of Fujiwara's theorems, therefore, establish a link between the two apparently different kinds of solution methods for the root-separation problem: the classical methods of solution using quadratic forms (such as Fujiwara's), and the methods of solution via matrix equations.

2. INERTIA THEOREMS

In this section, we mention some inertia theorems which will be used later in our proofs of Theorem 1 and Theorem 2.

THEOREM 3 (Carlson and Schneider [2]). *Let A be an $n\times n$ complex matrix with $\delta(A)=0$, and let there exists a nonsingular Hermitian matrix H such that*

$$AH+HA^*=N,$$

where N is positive semidefinite. Then $\operatorname{In}(A)=\operatorname{In}(H)$.

THEOREM 4 (Taussky [14], Hill [7], Wimmer [16]). *Let A be an $n\times n$ complex matrix. Then there exists an $n\times n$ Hermitian matrix H for which*

$A^*HA - H$ is positive definite if and only if A has no eigenvalues of modulus one. If $A^*HA - H$ is positive definite, then A has $\pi(H)$ $(\nu(H))$ eigenvalues of modulus greater (less) than one.

Using Theorem 4, we prove in the following another inertia theorem analogous to that of Carlson and Schneider. The proof of this theorem was suggested by Carlson [4].

THEOREM 5. *Let A be an $n \times n$ complex matrix with no eigenvalues of modulus one, and let there exists a nonsingular $n \times n$ Hermitian matrix H for which $A^*HA - H$ is positive semidefinite. Then A has $\pi(H)$ $(\nu(H))$ eigenvalues of modulus greater (less) than one.*

Proof. Let π_1 (ν_1) be the number of eigenvalues of A of modulus greater (less) than one. Let H_0 be the Hermitian matrix whose existence is guaranteed by Theorem 4. For each $t > 0$, $A^*(H + tH_0)A - (H + tH_0) = (A^*HA - H) + t(A^*H_0A - H_0)$ is positive definite, so that $\pi(H + tH_0) = \pi_1$, $\nu(H + tH_0) = \nu_1$. By continuity, $\pi(H) \leqslant \pi_1, \nu(H) \leqslant \nu_1$. But, as H is nonsingular, $n = \pi(H) + \nu(H) \leqslant \pi_1 + \nu_1 = n$, and we must have $\pi(H) = \pi_1$ and $\nu(H) = \nu_1$. ∎

3. TWO LEMMAS

In this section we present two lemmas. Lemma 1 forms the main tool of our proofs.

LEMMA 1. *The Bezout matrix B defined by two polynomials $f(x)$ and $g(x)$ of the same degree n is such that*

$$BA = A^T B,$$

where A is the companion matrix associated with $f(x)$.

Proof. Let

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_1 & a_2 & & \cdots & a_n \end{bmatrix};$$

consider now the matrix equation

$$XA = A^T X.$$

Since A is nonderogatory, by a result of Taussky and Zassenhaus [13] every solution matrix X is symmetric. Let $x_1, x_2 \dots x_n$ be the n rows of the matrix X. Then the above matrix equation is completely equivalent to

$$x_1 A = a_1 x_n,$$

$$x_i A = x_{i-1} + a_i x_n, \qquad i = 2,3,\dots,n.$$

Eliminating $x_1,\dots,x_{n-1}$, we get

$$x_n f(A) = 0.$$

Since, by the Cayley-Hamilton theorem, $f(A)=0$, x_n can be chosen arbitrarily. Thus the rows $x_1,x_2,\dots,x_n$ of a symmetric solution X of the matrix equation $XA=A^TX$ are such that

(i) x_n can be chosen arbitrarily, and
(ii) $x_1,\dots,x_{n-1}$ satisfy the recursive relation

$$x_{i-1} = x_i A - a_i x_n, \qquad i = n, n-1,\dots,3,2. \tag{4}$$

Again, if $Z_1, Z_2,\dots,Z_n$ are the n rows of the Bezout matrix B, then Barnett [1] and (independently) the author [5] have shown that

$$Z_{i-1} = Z_i A - a_i Z_n, \qquad i = n, n-1,\dots,3,2.$$

Thus, the first $n-1$ rows of B satisfy the same recursive relation as do those of X in (4). Since x_n is arbitrary, and in particular can be taken as the last row of B, the lemma is proved. ■

LEMMA 2 (Barnett [1], Datta [5]). *The Bezout matrix B defined by the polynomials*

$$f(x) = x^n - a_n x^{n-1} - \dots - a_2 x - a_1$$

and $g(x)$, of the same degree, satisfies[1]

$$B = Ug(A)$$

[1]It is to be noted that according to Barnett [1], $B=-Ug(A)$. This is because the Bezout matrix B defined by Barnett has the generating function $K(f)=[f(x)g(y)-f(y)g(x)]/[y-x]$.

where A is the companion matrix associated with $f(x)$, and

$$U=\begin{bmatrix} -a_2 & -a_3 & \cdots & -a_{n-1} & -a_n & 1\\ -a_3 & -a_4 & \cdots & -a_n & 1 & 0\\ -a_4 & -a_5 & \cdots & 1 & 0 & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ -a_n & 1 & \cdots & 0 & 0 & 0\\ 1 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}. \tag{5}$$

4. PROOF OF THEOREM 1

From (2), we have

$$F = DB \tag{6}$$

where

$$D = \mathrm{dg}\big(1,-1,1,\dots,(-1)^{n-1}\big).$$

Consider now two cases:

Case 1. n is Odd

Since, by Lemma 1, $BA=A^TB$, and the diagonal matrix D is such that $D^{-1}=D$, we obtain

$$DFA - A^TDF = 0,$$

that is,

$$FA - DA^TDF = 0. \tag{7}$$

Now, it is easy to check that

$$-DA^TD = A^T - U_1,$$

where

$$U_1=\begin{bmatrix} 0 & 0 & \cdots & 0 & 2a_1\\ 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & \cdots & 0 & 2a_3\\ 0 & 0 & \cdots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & 0\\ 0 & 0 & \cdots & 0 & 2a_n \end{bmatrix}.$$

So from (7) we get

$$FA + (A^T - U_1)F = 0,$$

that is,

$$FA + A^TF = U_1F = W_1. \tag{8}$$

It is easy to verify that the last row f_n of F is the negative of the transpose of the last column of U_1, namely $-(2a_1, 0, 2a_3, 0, 2a_5, \dots, 0, 2a_n)$, whence

$$W_1 = U_1F = -\tilde{f}_n^T f_n = -\begin{bmatrix} 4a_1^2 & 0 & 4a_1a_3 & 0 & 4a_1a_5 & \cdots & 0 & 4a_1a_n \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 4a_1a_3 & 0 & 4a_3^2 & 0 & 4a_3a_5 & \cdots & 0 & 4a_3a_n \\ 0 & 0 & 0 & 0 & 0 & & 0 & 0 \\ 4a_1a_5 & 0 & 4a_3a_5 & 0 & 4a_5^2 & \cdots & 0 & 4a_5a_n \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 4a_1a_n & 0 & 4a_3a_n & 0 & 0 & \cdots & 0 & 4a_n^2 \end{bmatrix}.$$

Case 2. n is Even

In this case we have

$$-DA^TD = A^T - U_2,$$

where

$$U_2 = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 2a_2 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 2a_4 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 2a_n \end{bmatrix}.$$

From (7), we have again

$$FA + A^TF = U_2F = W_2. \tag{9}$$

The last row f_n of F in this case also being the negative of the transpose of the last column of U_2, namely,

$$f_n = -(0, 2a_2, 0, 2a_4, 0, \dots, 2a_n),$$

we have

$$W_2 - f_n^T f_n = -\begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 4a_2^2 & 0 & \cdots & 0 & 4a_2 a_n \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 4a_4 a_2 & 0 & \cdots & 0 & 4a_4 a_n \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 4a_2 a_n & 0 & \cdots & 0 & 4a_n^2 \end{bmatrix}.$$

Clearly, W_1 and W_2 are negative semidefinite matrices. Again, it follows from Lemma 2 and the relation (6) that

$$F = DUf^*(A).$$

Since F is symmetric, we have

$$F = F^T = f^*(A^T) U^T D,$$

whence the nonsingularity of the Fujiwara matrix F implies the nonsingularity of the polynomial matrix $f^*(A^T)$. Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the n eigenvalues of A [equivalently the zeros of $f(x)$]; then the eigenvalues of $f^*(A^T) = f^*(A^*)$ are $f^*(\bar{\lambda}_1), f^*(\bar{\lambda}_2), \dots, f^*(\bar{\lambda}_n)$. Since

$$f^*(\bar{\lambda}_i) = (-1)^n \prod_{j=1}^{n} (\bar{\lambda}_i + \lambda_j) \quad (\neq 0),$$

and F satisfies the matrix equation (8) or (9), according as n is odd or even, it follows from a result of Carlson and Loewy [3] that the rank of the matrix $[A, W_I]$ is equal to $\operatorname{rank}[A, W_2] = n$. This again implies that $\delta(A) = 0$ [15]. Part of (a) of Theorem 1 now follows from Theorem 3. The proof of part (b) is similar.

5. PROOF OF THEOREM 2

From (3) we have

$$F = BP, \tag{10}$$

where P is the permutation matrix

$$P = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{11}$$

Then, since $BA = A^TB$, we get $A^TF - FPAP = 0$. Multiplication by A on both side gives

$$A^TFA - FPAPA = 0. \tag{12}$$

Now,

$$PAPA = \begin{bmatrix} a_1^2 & a_n + a_1a_2 & a_{n-1} + a_1a_3 & \cdots & a_2 + a_1a_n \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots 1 & \end{bmatrix}$$

$$= I + U_3,$$

where

$$U_3 = \begin{bmatrix} a_1^2 - 1 & a_n + a_1a_2 & a_{n-1} + a_1a_3 & \cdots & a_2 + a_1a_n \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

and I is an identity matrix of order n.

So from (12) we have

$$A^TFA - F(I + U_3) = 0.$$

That is, $A^TFA - F = FU_3 = W_3$.

Since the elements of the last column of F in this case are just the negatives of the corresponding elements of the first row u_1 of U_3, it is easy to verify that

$W_3 = -u_1^T u_1 =$

$$\begin{bmatrix} -(1-a_1^2)^2 & (1-a_1^2)(a_1a_2+a_n) & \cdots & (1-a_1^2)(a_1a_n+a_2) \\ (1-a_1^2)(a_1a_2+a_n) & -(a_1a_2+a_n)^2 & \cdots & -(a_1a_2+a_n)(a_2+a_1a_n) \\ \vdots & \vdots & & \vdots \\ (1-a_1^2)(a_1a_n+a_2) & -(a_1a_2+a_n)(a_2+a_1a_n) & \cdots & -(a_1a_n+a_2)^2 \end{bmatrix}$$

The matrix W_3 is clearly negative semidefinite. Again, by Lemma 2 and the relation (10), we get

$$F = Uf^*(A)P,$$

where U and P are as defined in (5) and (11). Since F is symmetric,

$$F = F^T = Pf^*(A^T)U^T.$$

Thus, if F is nonsingular, then so is $f^*(A^T)$.

The eigenvalues of $f^*(A^T) = f^*(A^*)$ are $f^*(\bar{\lambda}_1), f^*(\bar{\lambda}_2), \ldots, f^*(\bar{\lambda}_n)$, and for all i,

$$f^*(\bar{\lambda}_i) = \prod_{j=1}^{n} (1 - \bar{\lambda}_i\lambda_j)$$

is nonzero. It therefore follows that $\bar{\lambda}_i\lambda_i \neq 1$ for any i, i.e., that $f(x)$ has no eigenvalues of modulus one. Part (a) now follows from Theorem 5. The proof of part (b) in both directions is similar and will be omitted.

I would like to express my gratitude to Professor David Carlson of Oregon State University, Corvallis, Oregon, for making some valuable suggestions to improve the quality of the paper. I would also like to express my sincere thanks to Dr. S. Barnett of the University of Bradford, England for his interest in this work. I benefited greatly from correspondence with him.

REFERENCES

1 S. Barnett, A note on the Bezoutian matrix, *SIAM J. Appl. Math.* 22: 84–86 (1972).
2 D. Carlson and H. Schneider, Inertia theorems for matrices: the semidefinite case, *J. Math. Anal. Appl.* 6: 430–446 (1963).
3 D. Carlson and Raphael Loewy, On ranges of Lyapunov transformations, *Linear Algebra and Appl.* 8: 237–248 (1974).
4 D. Carlson, private communication.
5 B. N. Datta, Quadratic forms, matrix equations and the matrix eigenvalue problem, Ph.D. Thesis, University of Ottawa, Ottawa, Canada, March 1972.
6 M. Fujiwara, On algebraic equations whose roots lie in a circle or in a half plane (in German), *Math. Z.* 24: 161–169 (1926).
7 R. D. Hill, Inertia theory for simultaneously trianguable complex matrices, *Linear Algebra and Appl.* 2: 131–142 (1969).
8 J. L. Howland, The method of quadratic forms in dynamic stability problems, Third Canadian Conference in Applied Mechanics, Calgary, May 1971.
9 R. E. Kalman, On the Hermite-Fujiwara theorem in stability theory, *Quart. Appl. Math.*, pp. 279–282 (1965).
10 M. G. Krein and M. A. Naimark, *The Method of Symmetric and Hermitian Forms in the Theory of Separation of the Roots of Algebraic Equations* (in Russian), GNTI, Kharkov, 1936.
11 S. H. Lehnigk, Lyapunov's direct method and the number of zeros with positive real parts of a polynomial with constant complex coefficients, *SIAM J. Control* 5: 234–241 (1967).
12. P. C. Parks, A new proof of the Routh-Hurwitz stability criterion using the second method of Lyapunov, *Proc. Cambridge Philos. Soc.* 58: 694–702 (1962).
13 O. Taussky and H. Zassenhaus, On the similarity transformation between a matrix and its transpose, *Pacific J. Math.* 9: 893–896 (1959).
14 O. Taussky, Matrices C with $C^n \to 0$, *J. Algebra* 1: 5–10 (1964).
15 H. K. Wimmer, Inertia theorems for matrices, controllability and linear vibrations, *Linear Algebra and Appl.* 8: 337–344 (1974).
16 H. K. Wimmer, On the Ostrowski-Schneider inertia theorem, *J. Math. Anal. Appl.* 14: 164–169 (1973).

Received 12 March 1974; revised June 1977